

# DIGITIZATION OF KNOWLEDGE DOCUMENTS

Ms. Rekha Joshi

Librarian

A.D Institute of BCA, BCS, Khultabad, Aurangabad, Maharashtra

**Abstract:** The most important challenge in arrangement and execution of a digitization project relate to mechanic call imitations, budgetary constraint, copyright consideration, lack of rule guiding principle and lastly, the collection of materials for digitization. This paper studies the digitization initiatives of Digital Library of India, National Library. It focuses on Digitized Knowledge Documents in India like total number of digitized books, total number of digitized pages. The study also highlights the Center of Digitization and Digital Library of India, National Library, Krishikosh, Digitization, E-Granth Project. Digitization a archive in print or other physical media makes the record more valuable as well as more available. It is conceivable for a user to conduct a full-text search on a archive that's digitized. It is conceivable to make hyperlinks to lead a peruse to related things inside the content itself as well as to outside assets. Eventually, digitization does not cruel supplanting the conventional library collections and administrations; or maybe, it serves to improve them. A report can be changed over into advanced organize depending on the objective of digitization, conclusion client, accessibility of funds, etc.

**Keywords:** Digitization, DLI, National Library, Krishikosh,

**Introduction:** We have been utilizing printed report sources since a long time which are made accessible to us by the proficient endeavors of distributors, custodians and data researchers. But, with the diagram of data innovation different stages have been taken to diminish the endeavors by giving the data in a brief period of time. One of the critical applications of Data Innovation is the digitization of information record; the printed data alter to the advanced shape. The Digitization has changed the total situation of data in world. Today's era is advanced gadget society. Concurring to Pearce-Moses, Digitization is the method of changing analog report into twofold electronic/digital frame, particularly for capacity in difficult plate and utilize on computer. Digitization changes materials from analog groups that can be studied by individuals to a computerized arrange that can be studied as it were by machines. The gadgets like scanner or camera and altering program conjointly other gadgets can be utilized to digitize information substance. These advances permit the digitization of nearly imperative sorts of materials, Ph.D. Proposition, uncommon Books, compositions, photos, sound recordings, and pictures. Data is made in different designs i.e. Digitization is easy access to data assets. Digitized archive permits peruses to seek for collections quick and included all from anytime and anyplace. The method of digitization makes the imperceptible to be unmistakable. A number of users can get to the same archive at the same time without prevention. It also removes the inconvenience of separate, as clients don't have to travel to areas that have the difficult duplicates of materials. Digitization may be a time devouring and exceptionally costly wander, but it could be a brawny way to manage up with the issues of deficiency of critical periodicals, uncommon book, compositions and other writing in information assets centers within the creating world. Various organizations and Library are taking activities in digitizing their records, files of daily papers, uncommon books, compositions, artifacts, Ph.D. theses and M.Phil. Dissertation and other authentic reports and pictures. This makes a difference researchers, chairmen, peruses, and other data searchers to have, wide get to to advancements conceivable at a redress time which are prior exterior their space.

Digitization is characterized as, "Transformation of analog things into computerized organize for the reason of amplifying past get to and where suitability, to help with conservation, development. It is connected to all aspects of administrations given by the library." In basic words, digitization implies obtaining, checking, altering, changing, putting away and holding data in standardized and organized way with innovation back. With specialized scanners, application, reports are changing over into computerized designs and put away for advance reference by means of difficult circles, compact discs or web-based application'.

**Definition of Digitization:** Witten and David (2013) characterized Digitization as the method of taking conventional library materials that are in shape of books and papers, uncommon books to change over them to the electronic implies computerized frame where they can be put away and controlled by a computer. The US Founded of Historical center and Library Administrations characterizes digitization is the method of changing, making, and keeping up books, compositions, craftsmanship works, authentic records, photographs, diary, etc. in electronic representation so they can be seen through computers and savvy phones, portable workstation.

**1. Digital Library of India (DLI):** The Computerized collection of unreservedly open uncommon books collected from different libraries in India. DLI venture begun in early 2000 with the vision to file all the critical scholarly, creative and logical works of mankind and to protect carefully and make them accessible unreservedly for every one over Web for instruction, ponder, appreciation and for future eras. As a to begin with step in realizing this vision, it is proposed to form the Computerized Library with a free-to-read, searchable collection of one million books, overwhelmingly in Indian dialects. The Venture was started by the Office of the Foremost Logical Advisor to the Government of India and in this way taken over by the Department of Hardware and Data Innovation, Service of Communications and Data Innovation, Govt. of India. The thought was moreover to make a test bed for analysts to move forward filtering strategies, optical character acknowledgment, brilliantly ordering and in common to advance Indian Dialect Innovation Inquire about Digital Library of India has as of now 550,585 books with 191,657,791 pages (191.632 Million approx.) in Versatile Report Organize. This extend is financed by the Division of Gadgets and Data Innovation, the Service of Communication and Data Innovation, Govt. of India.

One of the objectives of the Digital Library of India is to supply bolster for full text indexing and looking based on OCR (optical character acknowledgment) advances available. The availability of online search allows clients to find significant data rapidly and dependably hence upgrading clients and investigate researcher victory in their inquire about endeavors.

Copyright act 1957 is The Indian copyright approach. On the Computerized Library of India site there accessible add up to books are copyright. All Digitized information archive are Noncommercial Employments As it were, The Advanced Library of India and its substance may as it were be utilized by enlisted clients for those noncommercial academic and instructive employments set forward in this. On the off chance that no anybody concur not to adjust, cloud, alter or expel any copyright take note or other attribution included within the Computerized Library, and not to authorize others to do so, Contact address is The DLI group, Supercomputer Instruction and Investigate Middle, Indian Established of Science, Bangalore, India. Digital Library of India work under in Indian Institute of Science, Bangalore in co-operation with MCIT, ERNET, NSF, CMU for the Government of India and twenty-one major participation centers. Table 1 show that total digitized collection has 5, 50,589 and total number of pages has 19, 16, 81,695 of digitized collection in available on Digital Library of India. Digitized books have maximum collection 1, 01,512 of C-DAC, Noida Institute Uttarpradesh. Number of pages maximum digitized has 3, 76, 84, 591 in Banasthali University, Niwai, Rajasthan. Table 2 show that Language wise digitized books with number of pages. English language Books has maximum 2,89,361 and number of pages has 11,59,32,078 (<http://www.dli.ernet.in/>) Data Access from 2017.

**Table 1: Institute wise list of Digitized documents with number of pages**

<b>Center wise Digitization Report</b>			
<b>Sr. no.</b>	<b>Scanning Center</b>	<b>Number of Books</b>	<b>Number of Pages</b>
1	AKCE, Universal Digital Library	7063	1026507
2	Allama Iqbal Library, University of Kashmir	17076	6102651
3	Anna University	231	106380
4	ASR, Melkote	1293	393852
5	Banasthali University, Rajasthan	94022	37684591
6	Bharatiya Gyanpeeth, Delhi	18	7612
7	Bharatiya Jnanpith, Delhi	292	92357
8	C-DAC, Noida	101512	31529575
9	C-DAK, Kolkata	70692	28929111
10	Central Library, University of Delhi	52	16537
11	City Central Library, Hyderabad	17257	3759927
12	Digital Library, Mulund	492	160184
13	Dr. Zakir Husain Library Jamia Millia Islamia	5	633
14	Durga Shah Municipal Library, Nainital	465	159370
15	Goa University	1611	191575
16	Government Central Press, Mumbai	165	60743
17	Gujrat Vidyapith Library	1	43
18	Gurukul Kangri Vishwavidyalaya, Haridwar	28	5200
19	IGNCA, Delhi	10195	4205808
20	IIAp, Bangalore	1626	586917
21	IIT, Allahabad	63741	20546413
22	IIT, Hyderabad	2240	714886
23	IISc, Bangalore	4422	1618253
24	Kerala Sahitya Akademi, Kerala	999	
25	Maulana Azad Library, A.M.U	6	1832
26	Namgyal Institute of Tibetology, Gangtok	2	1802
27	NIT, Sikkim	1015	1666183
28	North Eastern States Libraries	30156	9595316
29	Osmania University	7511	2660451
30	Pune University	1681	496638
31	Rashtrapati Bhavan Library	4831	2022258
32	RMSC, IIT-H	54995	18947375
33	Salar Jung Museum	6127	2338824
34	SASTRA University, Thanjavur	1867	575329
35	Sringeri Sharada Peetham	900	283292

36	State Central Library, Hyderabad	8802	3080062
37	S.V. Digital Library, Tirupati	15098	3860443
38	The Record And PersionOffice , War Department	4	280
39	University of Hyderabad	75	35140
40	University of Kashmir	76	19725
41	University of Pune	4	854
42	Unknown	11782	4163481
43	UOD, Delhi	6373	2622432
44	U.S. Department Of Agriculture	2	303
45	U.S. Department Of Education	1	1074
46	Veer Sewa Mandir, Dariya Ganj, Delhi	3783	1409476
	<b>Total :</b>	<b>550589</b>	<b>191681695</b>

(Source-<http://www.dli.ernet.in/>) Data Access in the year of 2017.

**Table 2: Language wise digitized books list with Number of pages**

<b>Language wise Digitization Report</b>			
<b>Sr. No.</b>	<b>Language</b>	<b>Number of Books</b>	<b>Number of Pages</b>
1	English	289361	115932078
2	Hindi	51724	15969189
3	Kannada	3336	720128
4	Arabic	7630	2484155
5	Persian	7624	2685979
6	Others	208	91774
7	Urdu	32265	9913618
8	Kashmiri	479	58146
9	Punjabi	402	93187
10	Sanskrit	35104	13884970
11	Gujarati	39817	6777655
12	Farsi	4	4644
13	Awadhi	2	323
14	Bhojpuri	1	160
15	Bengali	29668	8954446
16	Braj	5	2818
17	Danish	2	1062
18	Dutch	36	15024
19	French	1261	564854
20	Garo	116	13746
21	German	772	349179
22	Greek	23	11502

23	Irish	2	1226
24	Italian	69	39648
25	Khasi	600	62512
26	Kokborma	1	28
27	Kokborok	14	3258
28	Malayalam	582	16117
29	Manipuri	2714	373734
30	Marathi	3907	1044999
31	Mizo	280	47528
32	Nagari	1	64
33	Nepali	28	4952
34	Norwigen	3	882
35	Oriya	49	18454
36	Pali	94	32696
37	Pnar	4	404
38	Polish	5	3508
39	Portuguese	1450	168496
40	Prakrit	2	268
41	Rajasthani	4	756
42	Roman	2	506
43	Russian	8	2555
44	Sindhi	86	32819
45	Spanish	25	12605
46	Swedish	13	6161
47	Tamil	5376	1287932
48	Telugu	23313	4480770
49	Tibetan	1037	1674473
50	Turkish	133	25004
51	Multilingual	1931	495542
52	Unknown	6846	2753238
53	Assamese	2168	537824
54	Mongolian	2	352
55	Nagamese	1	27
	<b>Total</b>	<b>550590</b>	<b>191657975</b>

(Source-<http://www.dli.ernet.in/>) Data Access in the year of 2017.

The scanning and archiving of rare and brittle books and other documents are under way. English books and documents published before 1900 and Indian publications of pre 1920 are considered for digitization. So far 9140 selected books in Indian and English languages have already been scanned a total of over 3,20,000 pages. (<http://www.nationallibrary.gov.in/>)

## 2. National Library, Ministry of culture, Government of India: Digitization: Krishikosh - An Institutional Repository of India National :

### Agricultural Research System:

The KrishiKosh is an Regulation Store created beneath the beneath National Rural Inquire about Framework (NARS). In this extend, the mental yield of Indian NARS in different record shapes is captured, protected, filed to empower analysts to get to the substance online. It could be a one of a kind store of Information in horticulture and associated sciences, having a collection of theses, uncommon and important books, organization distributions, specialized bulletins, venture reports, addresses, preprints, reprints, field records and other archives accessible in several libraries of Investigate Organizing and State Rural Colleges spread all over the nation. KrishiKosh acts as a computerized stage to protect the institution's mental resources and to too oversee its utilize through open get to mandate.

Indian National Rural Inquire about Framework the Indian National Rural Inquire about Framework (NARS) may be a comprehensive data framework for agribusiness and associated sciences comprising of endless collection of uncommon and important books, reports, theses, studies, measurable information, maps, bulletins, bulletins, diary back volumes and other documents held by libraries of diverse Rural Investigate Organizing and State Rural Colleges spread all over the nation. In arrange to supply online get to to this profitable asset to analysts and researchers all over the world, an Regulation Store named 'KrishiKosh' has been made beneath the National Agrarian Advancement Extend (NAIP).

**Digitization:** The work of digitization of the inquire about records was taken up by four diverse centers, to be specific - Acharya N G Ranga Agricultural University, Hyderabad; Indian Rural Inquire about Established, Modern Delhi; Indian Veterinary Investigate Organized, Izatnagar; and College of Agrarian Sciences, Bengaluru. Using Open-Source program, the fabric within the store was organized into Institution astute collection of reports such as books, diaries, theses, reports, articles etc. The method of digitization counting checking, cleaning, trimming, altering transformation to text (OCR), change to PDF or JPG organize, is right now being done at four diverse Centre's. So distant, table 3.7 appear that more than 8.6 million pages from 19,495 archives have been digitized along side the fundamental Metadata.

**Digitization Centers:** Initially the repository has been developed at the following four libraries:

- i. Indian Agricultural Research Institute, New Delhi
- ii. University of Agricultural Sciences, Bangalore
- iii. Indian Veterinary Research Institute, Izatnagar
- iv. Acharya N G Ranga Agricultural University, Hyderabad

**Table 3: Digitized document in Krishikosh**

Sr. No.	Digitized Document	Quantity
1	8.6 million Pages digitized	19,495

(Source- <http://lib.hku.hk/etd2013/presentation>) Data Access in the year of 2017.

**Technology-KrishiKosh:** Applying Computerized Library benchmarks, a web interface has been created where client can browse through Title, Subject, put of distribution, Creator, subject key, lesson number, increase number, etc. Progressed look choice is additionally given counting confined get to to copyright things. Client can effortlessly enroll and subscribe for any collection and frequently get updates, etc. The 'KrishiKosh' could be a collaborative venture which imagines propelling an Indian rural information entrance over the Web and sharing its inquire about data counting Electronic Theses & Theses with the rest of the world utilizing most recent Data and Communication Advances.

**E-GRANTH extend:** The extend 'Strengthening of Advanced Library and Data Administration beneath (e-Granth Project)' has been approved by the National Rural Development Extend (NAIP)' which may be a World Bank financed activity beneath ICAR. Under this extend, a central foundation for all the exercises might be made at Library Modern Delhi. Within the to begin with stage of capacity building, twelve libraries have been chosen. After investigating the existing foundation and labor of these libraries, the hole will be filled up within the proposed arrange, so that they can meet the data request from NARS.

**Conclusion:** Whereas the targets of digitization activities contrast from organization to organization, the essential objective is to move forward the get to. Other goals incorporate taken a toll investment funds, conservation, keeping pace with innovation and data sharing. The foremost noteworthy challenges in arranging and executing of a digitization venture, relate to specialized restrictions, budgetary limitations, copyright contemplations, need of arrangement rules and finally, the determination of materials for digitization. Whereas unused and rising advances permit advanced data to be displayed in imaginative ways, the larger part of potential clients are improbable to have get to advanced equipment and program. Sharing of data among different educate is frequently confined by the utilize of incongruent program. Vital issues and issues relate to the choice prepare, arrangement of the materials, collaboration with the distributing division, the details for the digitization itself, investigate into improvement of optical character acknowledgment, investigate into a few record designator educe the taken a toll of capacity, programmed quality control instruments, modern language-based strategies for look and recovery, the advanced conservation of the records and the specialized organization to bolster all the regards. Attheturn of the century a move in accentuations is happened in digitization exercises. Libraries moved from digitizing highlights to digitizing total collections. Digitization advancements got to be bigger and so venture organization got to be a more vital issue. Noida Institute Uttarpradesh. Number of pages maximum digitized has 3, 76, 84, 591 in Banasthali University, Niwai, Rajasthan. Table 2 show that Language wise digitized books with number of pages. English language Books has maximum 2,89,361 and number of pages has 11,59,32,078. Digitized documents in Krishikosh 8.6 million Pages digitized. The scanning and archiving or rare and brittle books and other documents are under way. English books and documents published before 1900 and Indian publications of pre 1920 are considered for digitization. So far 9140 selected books in Indian and English languages have already been scanned a total of over 3,20,000 pages.

### References:

1. Alhaji, ibrahimusman. (2005). Digitization of library resources and the formation of Digital Libraries: a practical method. Retrieved from: <http://www.ais.up.ac>
2. Beatrice A. F. and Matthew Paris (2006). Digitization of Library Resources: Challenges and Implications for Planning and Policy. African and African American Studies International Journal, Vol. V, No. 22, Jul2006, Received from- <http://www.researchgate.in>
3. Borde S. A. (2017) Economic impact of Digital Resources and Digitization among University library, Ph. D. Thesis, Submitted by Dr.Babasaheb Ambedkar Marathwada University, Aurgabad, Maharashtra.
4. Bist, R. S. (2006). Digitization at Gandhi Smriti Library of LBSNAA: An Initiative to Preserve the Rare Books. In 4th Convention PLANNER- 2006. Mizoram University, Aizwal, November 9-10, 2006.
5. Dhule Subhash (2018). Digitization of libraries: Benefits and Challenges, *Knowledge librarian An International peer reviewed bilingual E-Journal of library and information science*, E-ISSN No. 2394-2479, <http://www.klibjlis.com>
6. Digital Library Federation. (2001). Registry of Digitized Books and Serial Publication. Retrieved from <http://www.digilib.org/collections>.
7. Gbaje S. (2007). Digitization and its challenges: Digital records and archival management workshop for members, federal capital territory, Abuja, Archives history Bureau committee and liaison officers, Arena House Kaduna.
8. Hannibal, Travis (2010). Estimating the Economic Impact of Mass Digitization Projects on Copyright Holders: Evidence from the Google Book Search Litigation Florida International University, Receive from- <http://papers.ssrn.com>
9. Khan Saima, shazia khan and mohsinaaftab (2015) Digitization and its impact on economy, international journal of digital library services, vol.5,april-june2015, issue-2, page no. 138-149, receive from- [http://www.Ijodls.in/uploads/3/6/0/3/3603729/vol-5\\_issue-2](http://www.Ijodls.in/uploads/3/6/0/3/3603729/vol-5_issue-2).
10. Kumar Rakesh and Kamble, V.T. (2015). Role in Digitization of Radio Libraries for Posterity and Access to the Civilization in India. E-Library Science Research Journal, Vol No.-4, Issue-4, Feb-2015, ISSN31984366, Received from- <http://www.Lsrj.in>
11. Lalitha, P. (2009). Importance of digitization of cultural and heritage materials. SRELS Journal of Information Management, 46(3).
12. Moses Weisser, C., & Walker, J. R. (1997). Electronic Theses and Dissertations: Digitizing Scholarship for Its Own Sake. The Journal of Electronic Publishing, 3(2). Retrieved from- <http://www.press.umich.edu/iep/03-02/etd.html>

13. Otubellu B N & Ume, L E. (2015). Digitization of library resources in academic libraries: challenges and implication. Retrieved from [www.iosrjournals.org](http://www.iosrjournals.org)
14. Pandey Prabhat. (2014). Digitization of Library Materials in Academic Libraries: Issues and Challenges. *Journal of Industrial and Intelligent Information* 2(2).
15. Pandey Prabhat and Misra Roli (2014). Digitization of knowledge document in Academic Libraries: Challenges and Issues. *Journal of Industrial and Intelligent Information*, Vol. 2, No.22, June2014, received from:[Http://www.jiii.org](http://www.jiii.org)
16. Powell, Christina Kelleher (2008). OPAC integration in the era of mass digitization: the Books experience. *Library Hi Tech*. 2008, Vol. 26, Issue 1.
17. Saima Khan, Shazia Khan and Mohsina Aftab (2015) Digitization and its impact on economy, *International Journal of Digital Library Services*, Vol.5, April-June2015, Issue-2, , Receive from: [http://www.ijodls.in/uploads/3/6/0/3/3603729/vol-5\\_issue-2](http://www.ijodls.in/uploads/3/6/0/3/3603729/vol-5_issue-2).
18. Sharma, Neelam. (2010). Automation and digitization of university libraries a study of university libraries. (Ph.D.Thesis, Punjab University), Retrieved from- <http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/80092>
19. Sivankalai, s. And chellappandi, p. (2014). A study of engineering users' approach on Digitization in academic library. *E-library science research journal*, Vol.2, Issue.5, March2014, ISSN: 23198436, Received from- <http://www.lsrj.in/uploaded/articles/223>.
20. Singh, s. n. (2006). Digitization initiatives and university libraries in india. 4th Convention PLANNER -2006, Mizoram University, Aizawl, Nov., 2006, INFLIBNET Centre, Ahmedabad, Retrieved from:<http://ir.inflibnet.ac.in/bitstream/1944/1220/1/202-209>.
21. Veeranjanyulu K. (2013) Krishikosh : An Institutional Repository of India National Agricultural Research System,
22. Witten and David Zhou,Y.(2013). Are Your Digital Documents Web Friendly: Making Scanned Documents Web Accessible. *Information technology and Libraries*, 29(3). Received from- [http://lib.hku.hk/etd2013/presentation/Veeranjanyulu %20-%20 Krishikosh](http://lib.hku.hk/etd2013/presentation/Veeranjanyulu%20-%20Krishikosh).
23. <http://www.dli.ernet.in/>
24. <http://www.nationallibrary.gov.in/>
25. <http://www.lib.hku.hk/etd2013/presentation>

